Downstream 28S Gene Sequences on the RNA Template Affect the Choice of Primer and the Accuracy of Initiation by the R2 Reverse Transcriptase

DONGMEI D. LUAN AND THOMAS H. EICKBUSH*

Department of Biology, University of Rochester, Rochester, New York 14627

Received 31 January 1996/Returned for modification 14 March 1996/Accepted 4 June 1996

R2 non-long terminal repeat retrotransposable elements insert at a unique site in the 28S rRNA genes of insects. The protein encoded by the single open reading frame of R2 is capable of conducting the initial steps of its integration in vitro. The protein nicks the noncoding strand of the 28S target DNA (the strand which serves as a template for RNA synthesis) and uses the 3' hydroxyl group exposed by this nick to prime reverse transcription of the R2 RNA template. This target-primed reverse transcription (TPRT) reaction requires that the RNA template contains the 250-nucleotide 3' untranslated region of the R2 element. If this RNA template ends at the precise 3' end of the R2 element, then extra nucleotides, which we refer to as nontemplated nucleotides, are added to the target before cDNA synthesis. The presence of downstream 28S gene sequences on the RNA template reduces the total efficiency but eliminates these nontemplated additions, resulting in nearly 90% of all TPRT products reproducing the 3' junctions seen in vivo. Templates with 5 to 10 nucleotides of the 28S sequence are used most efficiently in this in vitro TPRT reaction. The requirement for downstream 28S rRNA sequences probably explains why the R2 elements of most insects differ from the majority of non-long terminal repeat retrotransposons in that they do not contain an A-rich repeat at their 3' junction with the target DNA. The presence of downstream sequences on these in vitro R2 templates also revealed that the R2 reverse transcriptase can prime cDNA synthesis by using the 3' end of another RNA molecule. This RNA-primed cDNA synthesis is not based on sequence complementarity between the RNA primer and the R2 template. The ability to use the 3' end of a noncomplementary RNA molecule has also been seen with the reverse transcriptase of the mitochondrial Mauriceville plasmid of *Neurospora crassa*.

Non-long terminal repeat (non-LTR) retrotransposable elements are an abundant family of mobile elements found in most eukaryotes. The distinguishing features of these elements are that they encode an open reading frame with homology to reverse transcriptases but are not flanked by the terminal repeats associated with all retroviruses and all retrotransposable elements of the copia/Ty1 and gypsy/Ty3 groups (2, 11, 34). Evidence that the non-LTR retrotransposons encode a functional reverse transcriptase has been obtained for the jockey element of *Drosophila melanogaster* (18), CRE1 of *Crithidia fasciculata* (13), L1 of humans (26), and R2 of *Bombyx mori* (25). Direct evidence that these elements utilize an RNA intermediate in their mobility has been obtained from experiments in which inserted intron sequences were found to be accurately removed during the formation of new copies of the I element of *D. melanogaster* (28) and the L1 element of mice (12, 20).

Knowledge of the retrotransposition mechanism used by the non-LTR retrotransposable elements is still limited. Current models suggest that non-LTR elements retrotranspose by using the 3'-hydroxyl group at a DNA break to prime reverse transcription of their RNA transcripts (3, 11, 29). The only direct evidence for this model has been obtained from the R2 element of *B. mori* (25, 36). The single open reading frame of this R2 element was expressed in *Escherichia coli*, and the 120-kDa protein was found to have both endonuclease and reverse transcriptase activities. The R2 protein was shown to

first nick the noncoding strand (i.e., the strand used as a template for transcription) of the 28S gene at the insertion site, use the 3'-hydroxyl group exposed at the nick to prime reverse transcription of an R2 transcript, and then cleave the coding strand of the target DNA. This R2 retrotransposition mechanism has been termed target-primed reverse transcription (TPRT) (24, 25).

Evidence for the generality of the TPRT reaction can be found in recent studies of the mobilities of yeast mitochondrial group II introns. The mobile group II intron, *al2*, contains an open reading frame which encodes a reverse transcriptase activity (21). An in vitro assay used to detect the retrotransposition of *al2* indicated that an intron-encoded ribonucleoprotein complex was capable of cleaving the target DNA and using the 3' hydroxyl group of the noncoding strand to prime cDNA synthesis (39). Cleavage of the coding strand, which precedes the reverse transcription step, is catalyzed by the intron RNA sequences (40). The finding that both a group II intron and a non-LTR retrotransposable element utilize the same mechanism to prime reverse transcription suggests that this TPRT mechanism is widely used by members of each group of retroelements. Several arguments (27), including the phylogenetic analysis of the reverse transcriptase sequences from all retroelements (11, 38), suggest that TPRT is likely to predate the retroviral mechanism of using tRNA to prime reverse transcription (35).

In a previous report we demonstrated that the only R2 RNA sequences required for the TPRT reaction in vitro were those of the 250-nucleotide (nt) 3' untranslated region (3' UTR) of the R2 element (24). The most efficiently utilized RNA templates were those that ended at the precise R2-28S gene junction. However, few of the integrated cDNA products derived

* Corresponding author. Mailing address: Department of Biology, Hutchison Hall, University of Rochester, Rochester, NY 14627. Phone: (716) 275-7247. Fax: (716) 275-2070. Electronic mail address: teickbush@the.biology.rochester.edu.

from these templates accurately reproduced the 3' junctions of endogenous R2 elements of *B. mori*. Most of the in vitro integration products contained additional ("nontemplated") nucleotides inserted between the R2 sequences and the 28S gene target (24). In this study we have further analyzed the RNA requirements for accurate initiation of reverse transcription by testing R2 templates with sequences extending beyond the R2-28S junction. While significantly fewer integrations were observed, approximately 90% of the TPRT products generated with these templates had initiated reverse transcription at the first nucleotide of the R2 element, thus reproducing the junctions seen in vivo. The decrease in TPRT products was a result of the 3' end of another RNA molecule priming reverse transcription, thereby generating a cDNA product that was not integrated into the target site.

MATERIALS AND METHODS

Protein purification and TPRT assays. The R2 protein was purified from *E. coli* JM109/pR260 as previously described (24, 25) except that fractions from the Sepharose Q column were dialyzed against buffer containing 0.4 M NaCl and then loaded onto a native DNA-cellulose column that had been prewashed with the same buffer. All TPRT assays were performed in 200 mM NaCl–10 mM MgCl₂–50 mM Tris-HCl (pH 8.0)–5 mM dithiothreitol (DTT)–10 U of RNase inhibitor (Pharmacia)–25 μ M each deoxynucleoside triphosphate–125 nM α -³²S-dATP. Each 20- μ l TPRT reaction mixture contained 0.1 μ g of purified R2 protein, 0.2 μ g of in vitro-transcribed RNA, and 0.4 μ g of target DNA. Unless otherwise indicated, the pB109 target DNA (25) was predigested with *Hind*III and *Eco*RI. Products of the TPRT reactions were analyzed on 1% agarose gels. A relative quantification of the integration products was obtained by scanning autoradiographs of the gel (24). To determine the R2-28S junction sequences, the integration products were amplified by PCR and cloned into a modified M13 vector (6), and individual clones were sequenced.

DNA templates for in vitro transcription of RNAs. The plasmid construct used to generate the HR4/+50R RNA was generated by PCR amplification of clone B131 (4), using one oligonucleotide (5'-GCTGGAGAATTCGCGAG-3') complementary to sequences within the R2 element (starting at position 2186) and the second primer (5'-GATTCCTAGGGCAATTAGGTAAGTACGCGC-3') to prime synthesis on the 28S gene 50 bp downstream of the R2 insertion site. The PCR product was digested with *Hinc*II and *Bam*HI, and the resulting 0.85-kb fragment was cloned into the *Hinc*II and *Bam*HI sites of the vector pBSII(SK-) (Stratagene). This construct was predigested with *Bam*HI before being used as a template for in vitro transcription. Construct 249/+50R was made in a manner similar to that for HR4/+50R except that oligonucleotide 5'-TCAGTCGACG TGTGGTTGAGCCTTGC-3' was used as the R2 primer. HR4/+10V RNA and HR4/+28V RNA were made by in vitro transcription of the previously described HR4 construct (24) predigested with *Eco*RI and *Bam*HI, respectively. Construct pBmR2-249 (24) was predigested with *Eco*RI and used for the synthesis of 249/+10V RNA. DNA templates used for the synthesis of RNAs with different lengths of the 28S gene at the 3' end of the R2 element were generated by PCR amplification of the HR4/+50R construct by using the universal -40 primer complementary to the pBluescript vector upstream of T7 promoter region and the following primers corresponding to the individual 3' ends of RNA to be generated: 5'-TATTTTCGCCGATC-3' for HR4/+2R, 5'-GGCTATTTTCAT CGCG-3' for HR4/+5R, 5'-CATTTGGCTATTTTCATCG-3' for HR4/+10R, and 5'-CATTTGGCCGTTTTCATC-3' for HR4/+10R(GC). The DNA template for the 249/+10R RNA was generated by PCR amplification of the 249/+50R construct by using the universal primer and the +10R primer. All PCR products were treated with proteinase K (50 μ g/ml)–10 mM EDTA (pH 8.0), phenol-chloroform extracted, ethanol precipitated, and resuspended in TE (5 mM Tris-HCl [pH 8.0], 1 mM EDTA) before in vitro transcription. Conditions for the in vitro T7 transcription reactions and purification of the synthesized RNAs have been described previously (25).

Reverse transcription assay with the 3'-OH group on the target DNA blocked by dideoxadenylate (ddA). Ten micrograms of supercoiled pB109 DNA was nicked at the R2 target site by incubation with R2 protein for 10 min at 37°C in a reaction mixture containing 50 μ g of RNase A per ml, 200 mM NaCl, 10 mM MgCl₂, and 50 mM Tris-HCl (pH 8.0). The nicked pB109 DNA was treated with proteinase K (50 μ g/ml)–10 mM EDTA, phenol-chloroform extracted, ethanol precipitated, and resuspended in TE. Approximately 3 μ g of nicked pB109 was then incubated at 37°C for 45 min with 10 U of Klenow fragment (Bethesda Research Laboratories) in a 100- μ l reaction mixture containing 50 μ M each ddATP, dTTP, dGTP, and dCTP; 50 mM Tris-HCl (pH 8.0); 10 mM MgCl₂; and 50 mM NaCl. The reaction products were treated with proteinase K (50 μ g/ml)–10 mM EDTA, phenol-chloroform extracted, ethanol precipitated, and resuspended in TE. To ensure that only nicked, substituted DNA templates were used in the TPRT assay, the treated DNA was incubated with 3 U of T4 DNA ligase (Bethesda Research Laboratories) in a 50- μ l reaction mixture at room

temperature for 1 h under the conditions specified by the supplier. Following ligation, the reaction mixture was mixed with 3 μ l of loading dye (0.25% bromophenol blue and 40% sucrose in water) and run on a 0.7% agarose gel containing 1 μ g of ethidium bromide per ml. The nicked ddA-substituted open circular DNA was excised from the gel, purified with GeneClean (Bio 101), and used as a DNA substrate in the TPRT assay.

Determination of lengths of the nonintegrated reverse transcripts. A TPRT assay with a 10 $\times$ volume was conducted with supercoiled pB109 as the target DNA. After the incubation, the products were treated with proteinase K, phenol-chloroform extracted, ethanol precipitated, and resuspended in TE containing 5 mM DTT and 20 U of RNaseAguard. One half of the reaction product was treated with 50 μ g of RNase A per ml at 37°C for 30 min, while the second half remained untreated. Both samples were mixed with an equal volume of stop buffer (95% formamide, 20 mM EDTA, 0.05% bromophenol blue, 0.05% xylene cyanol), boiled for 5 min, and loaded on a 6% polyacrylamide–8 M urea gel. To accurately determine the site of initiation, the recovered cDNA was annealed to an oligonucleotide (5'-CCCCGAGCTCGTCC-3') which was complementary to a region containing an *Sst*I site 189 bp from the 3' end of the R2 element. The annealing reaction was carried out by boiling the cDNA with 0.1 pmol of the oligonucleotide for 90 s and then incubating for 10 min at 37°C. The reaction mixture was digested with *Sst*I in the presence of 50 μ g of RNase A per ml in a 10- μ l reaction mixture containing 50 mM Tris-HCl (pH 8.0), 10 mM MgCl₂, and 50 mM NaCl at 37°C for 1.5 h. After incubation, the reaction mixtures were mixed with 10 μ l of stop buffer, boiled for 5 min, and then loaded on a 6% polyacrylamide–8 M urea gel.

Junction sequences of the RNA-primed reverse transcription products. A TPRT assay with a 10 $\times$ volume was conducted with HR4/+10V RNA and pB109 DNA that had been predigested with *Eco*RI. The reaction products were treated with proteinase K, phenol-chloroform extracted, ethanol precipitated, and resuspended in TE containing 5 mM DTT and 20 U of RNaseAguard. The uninserted cDNA products were separated from the DNA substrate on a 1% agarose gel. The purified cDNA fragment was annealed to the primer R2-3'UTR, which is complementary to the cDNA strand 256 nt from the R2 3' end (25). Synthesis of the second cDNA strand with murine leukemia virus (MuLV) reverse transcriptase (Superscript RNase H⁻; Bethesda Research Laboratories) was carried out in a 20- μ l reaction mixture containing 50 mM Tris-HCl (pH 8.3), 75 mM KCl, 3 mM MgCl₂, 5 mM DTT, 40 U of RNaseAguard, and 5 mM each deoxynucleoside triphosphate at 37°C for 1 h. The reverse transcription products were incubated with 17 U of terminal transferase (U.S. Biochemicals) in the presence of 1 mM dCTP. The poly(C)-tailed DNA was diluted 50-fold and used as a template in an anchored PCR with the R2-3'UTR primer and the primer 5'-GGAATTCGCGATCCGGGGGGGGGGGGGG-3'. The PCR product was eluted from an agarose gel and cloned into the M13 vector for single-stranded sequencing of individual PCR products.

To determine if any exogenously added RNA could serve as a primer, a 57-nt vector RNA transcript (pBSK/Hd) was generated by T7 RNA polymerase transcription with *Hind*III-predigested pBSII(SK-) DNA. A TPRT reaction with a 10 $\times$ -volume was conducted as described above with 2 μ g of HR4/+10R RNA, 1 μ g of pB109 DNA cut with *Eco*RI, and 4 μ g of pBSK/Hd RNA. After purification of the nonintegrated cDNA and second-strand synthesis with MuLV reverse transcriptase, the products were directly amplified with an R2-3'UTR primer (25) and the primer 5'-CGAGGTGACGGTATCG-3', which is complementary to the pBSK/Hd RNA sequence starting 8 nt from its 3' end.

RESULTS

Effects of additional sequences at the 3' end of the R2 RNA template. The in vitro TPRT assay used in this study is summarized in Fig. 1. Sequences corresponding to the 250-bp 3'UTR of the R2 element are necessary and sufficient for the in vitro TPRT reaction. Longer R2 templates (~800 nt) are typically used in the reaction in order to increase the electrophoretic separation of the integration products (1.8 to 1.9 kb) from the target DNA (1.1 kb). RNA templates that end at the precise 3' junction of integrated R2 elements are used efficiently as templates in the TPRT reaction; however, in most of these priming events, extra nucleotides are added by the reverse transcriptase to the DNA target before the 3' end of the R2 template is engaged (24, 25). Because there is no obvious template for these extra sequences, we have referred to these extra sequences as nontemplated additions. These nontemplated additions result in over two-thirds of the in vitro TPRT products differing from the uniform sequence seen at the 3' end of all endogenous *B. mori* R2 elements (4).

To further investigate R2 protein-RNA interactions in the TPRT reaction, we have tested R2 RNA templates which con-

FIG. 1. Diagram of the in vitro assay used to study the TPRT reaction. The plasmid DNA substrate, pB109 (25), is predigested with *EcoRI* and *HindIII* to release the 1.1-kb 28S gene target DNA from the 2.7-kb pUC19 vector. Nicking of the noncoding strand of the target DNA and utilization of the 3'-OH to prime reverse transcription convert the 1.1-kb target DNA into a 1.9-kb product containing the cDNA-RNA heteroduplex. Cleavage of the second DNA strand only slightly increases the mobility of the TPRT product because the target site for the R2 insertion is only 60 bp from one end of the restriction fragment. The similarity in the mobilities of these two fragments allows the screening of a single band for quantification of TPRT products directly on an agarose gel or, if ^{35}S -dATP has been added to the reaction mixture, by autoradiography of the separated products. In this study all RNA templates tested contained the 798-nt 3' end of the R2 element and variable amounts of 28S rRNA or pBSII(KS-) vector sequences at the 3' end of the template.

tained various additional sequences located downstream (3') of the R2 element. Two types of downstream sequences are possible on the in vivo R2 RNA template. First, a noncanonical polyadenylation site may be present near the 3' end of the R2 element, which would allow RNA processing and the generation of a poly(A)-tailed transcript. Second, 28S rRNA sequences could be downstream of the R2 sequences if transcription of the R2 element involved neither transcription termination nor RNA processing at the 3' junction of the element. R2 RNAs with short poly(A) tails have already been tested and shown to serve as poor templates for the in vitro TPRT reaction (24). Polyadenylated R2 transcripts also generate an even greater range of nucleotide variation at the 3' junction, with less than 10% of the in vitro integration products reproducing the R2 3' junctions seen with endogenous *B. mori* R2 elements.

Figure 2 compares the efficiencies of various R2 templates with additional downstream sequences. The sequences of the various R2 RNA templates that have been tested are shown in Fig. 2A, and examples of the TPRT assays with different RNAs are shown in Fig. 2B. The templates tested included R2 RNAs with 2, 5, 10, or 50 nt of 28S gene sequences downstream of the R2 3' junction. To determine if specific sequences were being recognized, R2 transcripts containing 10 or 28 nt of the pBSII (SK-) polylinker sequences downstream of the R2 element were also tested. All templates with downstream sequences resulted in lower levels of TPRT products than the template without such additions (HR4). Of those RNAs with 28S gene sequences at the 3' end, the templates with 5 nt (HR4/+5R) or 10 nt (HR4/+10R) gave the highest level of TPRT product (approximately 50% of that of HR4 RNA). RNAs with either shorter (HR4/+2R) or longer (HR4/+50R) 28S sequences generated TPRT products at only 7 to 8% of the level generated by the HR4 template. Templates with downstream 28S gene sequences are more efficiently utilized than templates with downstream vector sequences: RNA template HR4/+10V was used at one-fourth of the efficiency of RNA template

A

RNA template 3' ends		Template efficiency
HR4	GAUCCGGCGAUGAAAA	1
HR4/8A	GAUCCGGCGAUGAAAAAAA	0.15 (0.04)
HR4/+2R	GAUCCGGCGAUGAAAAUA	0.08 (0.02)
HR4/+5R	GAUCCGGCGAUGAAAAUAGCC	0.52 (0.10)
HR4/+10R	GAUCCGGCGAUGAAAAUAGCCAAUG	0.56 (0.06)
HR4/+50R	GAUCCGGCGAUGAAAAUAGCCAAUG [+40 nt.]	0.07 (0.02)
HR4/+10V	GAUCCGGCGAUGAAAAUAUUCGAAUU	0.12 (0.06)
HR4/+28V	GAUCCGGCGAUGAAAAUAUUCGAAUU [+18 nt.]	0.04 (0.01)
HR4/+10R(GC)	GAUCCGGCGAUGAAAAUAGCCAAUG	0.06 (0.01)

R2 sequences Downstream sequences

B

FIG. 2. Effects of additional nucleotide sequences on the ability of HR4 RNA to support TPRT. (A) Summary of the RNA templates used and the efficiencies with which they served as templates in the TPRT reaction. The dashed vertical line represents the in vivo junction of the R2 element with the 28S gene. Nucleotide differences between the vector and 28S rRNA sequences are underlined. Template efficiencies are given as the fraction of 1.8- to 1.9-kb products generated relative to that generated by the HR4 template. These numbers are based on the averages of three or more experiments, with standard deviations shown in parentheses. (B) Ethidium bromide-stained 1% agarose gel of the TPRT reaction with the RNA templates shown in panel A. The sizes of the DNA length standards in the left lane are shown (in kilodaltons). The various substrates and products of the reaction are as follows: PUC, 2.7-kb pUC19 vector DNA; TPRT, 1.9-kb TPRT product; Target, 1.1-kb DNA fragment containing the 28S gene target; RPRT, 0.9-kb RNA-primed reverse transcription product (see Fig. 6 and 7).

HR4/+10R. Finally, in addition to the 1.8- to 1.9-kb TPRT product, all RNA templates with 3' extensions also generated a product band with an apparent length of approximately 0.9 kb. This lower band was not seen with the HR4 template. The origin of this new product band is discussed below.

Downstream RNA sequences result in precise initiation of reverse transcription at the R2 3' end. To determine where reverse transcription initiated on the R2 templates that contained either downstream 28S or vector sequences, the products of the in vitro reactions were PCR amplified, and individual products were cloned and sequenced (see Materials and Methods). The cDNA-28S gene junction sequences obtained from the various RNA templates are summarized in Fig. 3. With each RNA template most of the TPRT products resulted from reverse transcriptions that initiated at the R2-28S or R2-vector junction on the RNA template. The remaining products contained nontemplated nucleotides inserted between the R2 sequences and the target DNA (shaded nucleotides) or initiated reverse transcription at internal sites within the RNA template. Of the 49 clones sequenced from the RNA templates

			No. clones
HR4/+2R RNA	..TAGGCCGCTACTTTT	ATCGGTT..	17
	..TAGGCCGCTACTTTTCT	ATCGGTT..	1
HR4/+10R RNA	..TAGGCCGCTACTTTT	ATCGGTT..	16
	..TAGGCCGCTACTTTT	ATCGGTT..	1
	..TAGGCCGCTACCGTG	ATCGGTT..	1
	..GGGAATATATAAA ⁻¹⁵⁸	ATCGGTT..	1
HR4/+50R RNA	..TAGGCCGCTACTTTT	ATCGGTT..	10
	..TTTCGCCGCGCTA ⁻¹⁰⁸	ATCGGTT..	2
HR4/+10V RNA	..TAGGCCGCTACTTTT	ATCGGTT..	8
	..ACCTGAGTCACTA ⁻³³	ATCGGTT..	1
HR4/+28V RNA	..TAGGCCGCTACTTTT	ATCGGTT..	6
	..CGATCCTAATCC ⁻⁶²	ATCGGTT..	1
	..GGCCACGATCGG ⁻⁹⁵	ATCGGTT..	1
	..ACGCCACGATC ⁻⁹⁷	ATCGGTT..	1
	..CTTATATATTTT ⁻¹⁶²	ATCGGTT..	1
	..CTTGGCTCCCT ⁻¹⁶⁸	ATCGGTT..	1
HR4/+10R(GC) RNA	..TAGGCCGCTACTTTT	ATCGGTT..	13
	..TAGGCCGCTACTTTT	ATCGGTT..	1
	..TAGGCCGCTACTTTT	ATCGGTT..	1
	..TAGGCCGCTACTTTTATCGGAAAA	ATCGGTT..	1
	..TAGGCCGCTACTTTTCTACTTTT	ATCGGTT..	1

R2 template ←→ 28S gene

FIG. 3. 3' junctions obtained with the HR4 templates containing various additions at the 3' end. TPRT reactions similar to those for which results are shown in Fig. 2B were PCR amplified with one primer complementary to the R2 sequences and a second primer complementary to the 28S gene sequences downstream of the target site (see Materials and Methods). The PCR products were cloned, and individual products were sequenced. The sequences of the cDNA strands are shown. R2 sequences within the RNA template are shown to the left of the vertical line, and 28S gene sequences are to the right of the vertical line. Nucleotides present at the junction that would not be predicted on the basis of simple reverse transcription of the RNA template are shaded. Some of the sequenced products resulted from reverse transcription that initiated at internal positions within the R2 template. The distances of these internal initiation sites from the 3' ends of the R2 sequences are indicated. The number of clones corresponding to each sequence is indicated to the right. The sequences at the 3' end of each RNA template used in the reactions are given in Fig. 2A.

with downstream 28S sequences (combined data from the HR4/+2R, +10R, and +50R RNA templates), 43 (88%) reproduced the R2 junction seen in vivo. Additional nucleotides were added in only three instances, and four products resulted from internal initiations of reverse transcription (one of these clones also contained additional nucleotides). The presence of vector sequences downstream of the R2 sequences on the RNA template also allowed accurate initiation of reverse transcription at the R2-vector junction. Fourteen of 20 clones (70%) derived from HR4/+10V and HR4/+28V had initiated reverse transcription at this position (Fig. 3). The six clones that did not reproduce the in vivo junction resulted from internal initiations between 33 and 168 nt into the R2 element. These values for RNA templates containing downstream sequences contrast significantly with the TPRT products obtained with HR4 RNA (25). In the case of the HR4 template, only 4 of 11 clones (36%) sequenced from the reaction reproduced the in vivo R2 junctions. The remaining seven clones contained extra nucleotides added to the DNA target before the reverse transcriptase engaged the R2 template. Thus, the

presence of downstream 28S or vector sequences on the R2 RNA template increases the precision with which the R2 reverse transcriptase can initiate cDNA synthesis at the precise 3' junction of the R2 element and significantly decreases its tendency to add nontemplated nucleotides before engaging the RNA template.

What is responsible for this more accurate initiation of cDNA synthesis at the R2 3' junction? One possibility is that the target DNA downstream of the nicked site is denatured by the R2 protein, which allows the downstream 28S rRNA sequences on the template to base pair with the lower strand of the DNA target. The finding that downstream vector sequences also increase accurate initiation would seem to argue against this model. However, by coincidence the first two nucleotides of the downstream vector sequence, UA, are also the first two nucleotides of the 28S rRNA sequence (Fig. 2A). Thus, it was possible that the accurate initiation seen with both downstream 28S and vector sequences was due to hybridization of the first two nucleotides of the RNA to the target DNA. This model is also supported by the finding that template HR4/+2R with only these two downstream nucleotides can increase the accuracy of priming. We therefore generated another RNA template that contained 10 nt of downstream 28S rRNA, but the first two nucleotides of this sequence were changed to GC [HR4/+10R(GC)] (Fig. 2A). Such an RNA template would be unable to base pair with the first two nucleotides of the target DNA downstream of the nicked site. When HR4/+10R(GC) RNA was used as the template in the TPRT reaction, the template efficiency was reduced from 56 to 6% (Fig. 2A); however, the accuracy of the priming reaction remained high. All 17 of the clones recovered from the HR4/+10R(GC) RNA reaction derived from reverse transcription that initiated at the precise 3' R2-28S junction (Fig. 3). Four clones did not reproduce the correct 3' R2-28S junction because nontemplated nucleotides were added to the target DNA before the RNA template was engaged. Therefore, the ability of the R2 protein to find the precise R2 3' junction on RNA templates with downstream sequences does not require that the first two nucleotides of the downstream sequence be annealed to the target DNA. This experiment does not exclude the possibility that DNA sequences further downstream of the nicked site are annealed to the RNA template.

Downstream sequences on the R2 RNA template generated another reverse transcription product. As shown in Fig. 2B, all RNA templates containing additional downstream sequences resulted in the formation of a 0.9-kb band not seen with the HR4 template. Several experiments have demonstrated that this lower band represented a reverse transcript of the R2 RNA template in the reaction (data not shown). First, this band was labeled by ³⁵S-dATP added to the TPRT assays. Second, when R2 RNA templates of 249, 417, or 798 nt in length were used in the reaction, the size of the lower band was directly proportional to the size of the input RNA. Third, the band was sensitive to RNase H digestion. Finally, when an oligonucleotide complementary to the RNA synonymous strand of the R2 element was used as a sequencing primer, the purified 0.9-kb band was shown to contain a reverse transcript of the R2 template. Because the reverse transcription reaction which generated this lower band did not result in the attachment of the cDNA to the DNA target site, we refer to this band as a nonintegrated reverse transcript. Like the TPRT product, the nonintegrated product is produced only from R2 RNA transcripts and not from other RNAs (e.g., total *E. coli* RNA or vector RNA templates).

Because the nonintegrated products did not appear to use the 3'-hydroxyl group of the target DNA to prime the reaction,

FIG. 4. Synthesis of the nonintegrated reverse transcription product requires the 28S DNA target site. Shown is an ethidium bromide-stained agarose gel of the TPRT reaction with HR4 RNA and HR4/+10V RNA as templates. *EcoRI-HindIII*-predigested plasmid pB109 was used in lanes 1 and 2, while *EcoRI-HindIII*-predigested pUC19 DNA was used in lanes 3 and 4. DNA size standards and the various substrates and products of the reaction are the same as in Fig. 2. Because the reaction products were not treated with RNase A before being loaded on the agarose gel, the input RNA templates in each reaction can be seen as the bottom diffuse band in each lane (labeled RNA).

the first question to be addressed was whether formation of the 0.9-kb band would occur in the absence of the 28S gene target site. Shown in Fig. 4 are the results of TPRT experiments with HR4 and HR4/+10V RNA templates in the presence (pB109) and the absence (pUC19) of the 28S target site. In the presence of the target site, the HR4 RNA template gave rise to the 1.9-kb TPRT product, while the HR4/+10V RNA template generated both the 1.9-kb band (at reduced levels) and the 0.9-kb band. In the absence of the target site (pUC19), neither the 1.9-kb integrated product nor the 0.9-kb nonintegrated product was generated. Similar results have been obtained with HR4/+28V, HR4/+10R, and HR4/+50R (data not shown). These results indicate that like the TPRT reaction itself, the nonintegrated reverse-transcribed product requires the presence of the 28S target site.

The requirement for the 28S target site to generate the 0.9-kb product raised the possibility that this product was also derived from a target DNA-priming mechanism but that the RNA:DNA heteroduplexes were cleaved from the substrate DNA after synthesis. This putative cleavage activity could be viewed as similar to the original nicking activity of the endonuclease for the noncoding strand of the target DNA. To determine whether the 0.9-kb product utilized the nicked target DNA to prime reverse transcription, a DNA substrate for the reaction was generated such that the terminal 3' nucleotide at the nicked site (deoxyadenylate [dA]) was replaced by a dideoxynucleotide (ddA) (see Materials and Methods). This ddA substitution should enable the R2 protein to bind specifically to the target site, but it would eliminate the ability of the nick to be used as a primer for DNA polymerization.

The results of the TPRT assays with the ddA-substituted target DNA and both HR4 and HR4/+10V templates are shown in Fig. 5. It should be noted that the pB109 DNA used in these reactions was not predigested with *EcoRI* and *HindIII* in order that nonsubstituted DNA could be readily separated from the ddA-substituted DNA by treatment with T4 DNA ligase (see Materials and Methods). When both supercoiled

pB109 DNA (U) and prenicked, nonsubstituted pB109 DNA (dA) were used as substrates, the high-molecular-weight integrated TPRT products were generated. The nonintegrated products were observed only with the HR4/+10V RNA. When the ddA-substituted pB109 DNA was used as a substrate, generation of the TPRT product was completely eliminated for both the HR4 and HR4/+10V templates, while the 0.9-kb product was still generated with the latter template. Thus, the 3'-hydroxyl group released by the nick on the target DNA is not used to prime the reverse transcription that generates the 0.9-kb product. The finding that the DNA target is still required for this reaction, however, implies that the R2 protein is in the proper conformation to recognize or reverse transcribe the R2 template only after it has associated with the target site.

The 3' end of another RNA molecule primes formation of the nonintegrated reverse transcript. If the 0.9-kb reverse transcription product does not require the 3'-hydroxyl group of the target DNA to prime reverse transcription, then what is the primer? The reverse transcriptase of the Mauriceville plasmid from *Neurospora crassa* can initiate cDNA synthesis either de novo or primed by a 3'-OH from other nucleic acids (32, 33). De novo synthesis by the *N. crassa* enzyme occurs only at high nucleotide concentrations, because the K_m for dGTP, the first nucleotide in the de novo synthesis of the cDNA, is much higher for de novo synthesis than for priming by another nucleic acid (33). Assuming that the R2 enzyme would also show a higher K_m for the initiating nucleotide if it was capable of de novo synthesis, the ratio of integrated products (1.9 kb) to nonintegrated products (0.9 kb) was monitored when the deoxynucleoside triphosphate concentrations in the assay were varied from 5 to 200 μ M. No change in the ratio of integrated to nonintegrated products was detected (data not shown).

FIG. 5. Effects of a blocked 3'-hydroxyl group at the target site on the R2 reverse transcription reaction. (A) Agarose gel of the TPRT reaction with different forms of pB109 as target DNA. U, untreated DNA (supercoiled pB109); dA, pB109 DNA prenicked at the target site by the R2 protein; ddA, pB109 DNA nicked by the R2 protein with the deoxyadenylate residue at the 3' end of the nick replaced by a dideoxynucleotide. In lanes 1 to 3, HR4 RNA was used as the template; in lanes 4 to 6, HR4/+10V RNA was used as the template. (B) Autoradiograph of the gel in panel A. Because the target DNA in these reactions is a 3.7-kb circle, the TPRT products appear on the agarose gel as two high-molecular-weight products (labeled TPRT). The upper product has not undergone second-strand cleavage and thus is a circle with a branching cDNA component. The lower band has undergone second-strand cleavage resulting in a linear product. The nonintegrated reverse transcription product is labeled RPRT. The initial substrate DNA found in lanes 1 and 4 is in either the open circular (oc), linear (l), or supercoiled (sc) form. Because of the loss of DNA during the various manipulations required to modify the target site, less DNA is present in lanes 2, 3, 5, and 6 than in lanes 1 and 4.

FIG. 6. Length of the nonintegrated cDNA products generated in the TPRT assay. (A) The TPRT reaction was conducted with supercoiled pB109 DNA and the RNA template 249/+10V. The nonintegrated cDNA product was isolated from an agarose gel and analyzed on a 6% polyacrylamide-8 M urea gel. The presence or absence of RNase A treatment of the cDNA product before loading on the polyacrylamide gel is indicated above the lanes. The series of higher-molecular-weight bands remaining after RNase A treatment can be a result either of the R2 reverse transcriptase switching to another RNA template when the polymerase reaches the end of the initial R2 RNA template or of priming by the 3' end of a DNA molecule. (B) The nonintegrated cDNA product with the 249/+10V RNA was digested with *Sst*I in the presence of RNase A and loaded on a 6% polyacrylamide-8 M urea gel (lane 5). Lanes 1 to 4 correspond to a sequencing ladder of the 3' R2-28S gene junction starting at the site of the *Sst*I cleavage. The numbers to the right of lane 5 indicate nucleotide positions relative to the 3' end of the R2 element (position 0).

While indirect, this result suggested that the 0.9-kb fragment generated by the R2 reverse transcriptase was probably not a result of *de novo* synthesis.

To explore the possibility that the 3' end of another nucleic acid molecule was being used as a primer, the lengths of the nonintegrative reverse transcription products were determined on a 6% denaturing polyacrylamide gel with the shorter RNA template 249/+10V. This template contained the 249-bp 3' UTR of R2, 30 bp of vector sequences at its 5' end, and 10 bp of vector sequences at its 3' end. After the TPRT reaction, the low-molecular-weight nonintegrated cDNA product was recovered from an agarose gel and divided into two aliquots. One of the aliquots was treated with RNase A before being loaded onto the denaturing polyacrylamide gel. The second aliquot was not treated. The results of one such experiment are shown in Fig. 6A. In the absence of RNase A treatment, the cDNA from the nonintegrated reverse transcription appeared as several diffuse bands all greater than the 290 nt, the size expected from simple cDNA synthesis of the R2 RNA template. After RNase A treatment, the cDNA products migrated as a diffuse band approximately 290 nt in length. This suggests that most of the cDNA was primed by an RNA molecule and that initiation of reverse transcription occurred near the 3' end of the RNA template.

To more accurately determine the initiation sites for this RNA-primed reverse transcription, the experiment with the

results shown in Fig. 6A was modified to eliminate the heterogeneity at the 3' end of the cDNA due to premature terminations and the tendency of the R2 reverse transcriptase to template switch to another RNA (23). After the TPRT reaction, the nonintegrated cDNA fragment was denatured and annealed to an oligonucleotide complementary to a region containing an *Sst*I restriction site within the 3' UTR region of the R2 element. After annealing, the cDNA was cleaved with *Sst*I in the presence of RNase A to generate short cDNA fragments extending from the site of initiation to the internal *Sst*I site. The *Sst*I-digested product was analyzed by polyacrylamide gel electrophoresis adjacent to a DNA sequencing ladder starting at the site of *Sst*I cleavage within the R2 sequence. As shown in Fig. 6B, most of the reverse transcription initiated on the RNA template within a few nucleotides of the junction between the R2 and pBSK vector sequences. Initiations also occurred at the end of the RNA template (+10) and at internal sites within the R2 sequences (e.g., -4).

Evidence that priming of the nonintegrated cDNA by another RNA molecule involved a typical 3'-to-5' linkage was obtained by using a modified anchored PCR approach (22) as described in Fig. 7A. The nonintegrated cDNA band from an HR4/+10V RNA reaction was isolated and used as the template for synthesis of the second DNA strand by MuLV reverse transcriptase (see Materials and Methods). Any RNA (or DNA) fragment that was linked to the 5' end of the first cDNA strand by a normal 5'-to-3' linkage would also be used as a template by the MuLV reverse transcriptase. The 3' end of the second cDNA strand was then C tailed and PCR amplified. The junctions between the RNA primers and cDNA sequences of 12 separate clones from this reaction are shown in Fig. 7B. In 7 of the 12 junctions obtained by this method, the RNA used as the primer was the R2 RNA template itself (the sequence of the 3' end of the R2 template is shown at the top of Fig. 7B). Because the R2 RNA template is the most abundant RNA species present in the *in vitro* assay, it was not surprising that seven of the cDNA products were primed by another molecule of the RNA template itself. The remaining five junctions shown in Fig. 7B were primed by oligonucleotides of unknown origin. This priming is presumably by the wide range of short degradation products of *E. coli* RNA contaminating our R2 protein preparations. For all 12 sequenced products, cDNA synthesis initiated at sites near the R2-vector sequence junction of the RNA template, consistent with the data of Fig. 6B.

To determine if any RNA can function as a primer in the nonintegrative reverse transcription reaction, a 57-nt RNA molecule containing only vector sequences (pBSK/Hd) was added to the reverse transcription assay in a 30-fold molar excess with respect to the HR4/+10R RNA used as template. After reverse transcription, the nonintegrated cDNA band was again isolated and analyzed by PCR amplification with an R2-3'UTR primer and a primer complementary to the pBSK/Hd RNA (see Materials and Methods). As shown in Fig. 7C, all of the products obtained by this PCR amplification were primed by the 3' end of the pBSK/Hd RNA (one clone contained a 1-bp deletion from the 3' end of the primer RNA). Approximately half of the products contained an additional, GC-rich nucleotide sequence between the RNA primer and cDNA sequence. Reverse transcription of the R2 template began either at site +5 or +1 with respect to the R2 junction or at internal sites well within the R2 sequences. The combined results of Fig. 7B and C suggest that a variety of exogenously added RNA molecules are likely to serve as primers for reverse transcription of R2 templates containing downstream 28S or vector sequences.

We have also tested whether single-stranded or double-

FIG. 8. Model of the initiation of cDNA synthesis primed either by the nick on the target DNA (A) or by the 3' end of another RNA molecule (B). The 28S gene target site is drawn as a nicked and bent DNA molecule. Downstream 28S rRNA sequences on the R2 RNA template are indicated by the dashed box. (A) The 28S downstream RNA sequences assume an appropriate conformation which allows precise positioning of the R2 RNA template for initiation of DNA synthesis at the exact border of the R2 and 28S sequences. (B) The downstream rRNA sequences are in a conformation which prevents the proper positioning of the RNA template for priming by the 3'-OH of the DNA nick. As a result, the 3' end of another RNA molecule is used to prime DNA synthesis again near the R2-28S border on the RNA template. The RNA hairpin structures drawn for the downstream RNA sequences are for stylistic purposes to imply a precise structure and are not intended to imply that the specific structure of this sequence is known.

sequence to hybridize to the denatured DNA at the target site. Evidence against this model was obtained by using the RNA template HR4/+10R(GC). This template is identical to HR4/+10R except that the first two nucleotides of the 28S sequence, TA, were replaced by GC. These substitutions would eliminate the ability of the RNA template to hybridize to the first two downstream nucleotides of the target DNA at the nicked site. We found that while the efficiency of the reaction was reduced nearly 10-fold, all of the TPRT products obtained with the HR4/+10R(GC) template had still initiated reverse transcription at the precise R2-28S junction. This finding suggests that the 28S gene sequences on the template may contribute to the binding and correct positioning of the RNA template on the R2 protein rather than base pair to the target site. However, direct studies of the binding of the RNA templates to the R2 protein, as well as additional mutagenesis of the RNA template and DNA target, are needed to fully resolve this issue.

The finding that downstream 28S sequences of the R2 template enable the R2 reverse transcriptase to initiate cDNA synthesis at the correct site supports the model that the transcript used *in vivo* for R2 retrotransposition is not processed at the 3' junction with the 28S gene. While the templates that work best *in vitro* are those containing 5 to 10 nt of the 28S rRNA, such a template may not be the template used *in vivo*. First, there is no obvious mechanism for the processing of the RNA transcript a short distance downstream of the R2 3' junction. Second, we have found that downstream 28S rRNA sequences hundreds of nucleotides in length can sometimes be utilized in the TPRT reaction at efficiencies between those of HR4/+10R and HR4/+50R (23). The variability seen with these longer additions may result from the difficulty these longer 28S sequences have in assuming a specific higher-order structure in our *in vitro* reactions. Third, current models of the secondary structure of 28S rRNA (16) predict that the 80 nt downstream of the R2 insertion site base pair with segments of the 28S sequence approximately 400 bp upstream of the R2 site. Thus, the RNA templates that eventually work most efficiently and accurately may be those that contain the R2 elements as part of a total 28S rRNA transcript. Such a finding would strongly argue that R2 elements are cotranscribed with the 28S genes and do not contain their own 5' promoter.

The ability of the R2 reverse transcriptase to recognize the downstream 28S gene sequences may explain why R2 elements

differ from most other non-LTR retrotransposable elements in that they do not contain a simple A-rich repeat at their 3' junction with the DNA target site (10). Two arguments suggest that the reverse transcriptase of most non-LTR retrotransposons would not bind downstream sequences on their RNA templates. First, many non-LTR elements appear to contain polyadenylation sites which would eliminate the downstream sequences from the RNA template (2, 10). Second, most non-LTR elements without such polyadenylation sites insert at multiple locations in the genome and thus are unlikely to recognize downstream sequences on their templates because these downstream sequences would be different for each insertion. Without the ability to bind downstream sequences, the reverse transcriptase encoded by most non-LTR retrotransposable elements may not be able to accurately position the RNA template for cDNA synthesis. As a consequence, reverse transcription by these elements would be more similar to that for the R2 protein with the HR4 template, in that initiation could occur at multiple sites and involve the addition of extra, nontemplated nucleotides to the target DNA (24, 25), thereby generating the short A-rich repeats seen at the 3' ends of most elements (2, 10). Only non-LTR retrotransposable elements that insert at unique locations in the genome would be able to evolve the ability to bind downstream sequences on their RNA template. However, CRE1 (14), CRE2 (30), SLACS (1), and CZAR (31) elements, which insert specifically into the leader exon of trypanosomes, and Tx1, which inserts specifically into another mobile element in frogs (15), contain variable A-rich tails and thus presumably do not recognize downstream sequences on the RNA templates. Only R1 elements of insects and R4 elements of nematodes, elements which also insert at specific sites in the 28S gene (6, 19), do not contain A-rich 3' junctions and thus may also utilize recognition of downstream 28S gene sequences to accurately initiate reverse transcription. Even in the cases of R1, R2, and R4, some species contain elements that contain A-rich 3' tails (9, 19), suggesting that recognition of the downstream 28S sequences may be readily lost by the reverse transcriptases of these elements.

The second major finding of this study is that all of the RNA templates tested which contained downstream sequences also generated a reverse-transcribed product that had been primed by the 3' end of another RNA molecule (Fig. 8B). Several RNAs were able to serve as this primer, suggesting there is little or no sequence requirement. The RNA-primed reverse transcription reaction, like the TPRT reaction itself, requires the 28S gene target site. In the absence of the DNA target, the R2 reverse transcriptase is similar to retroviral reverse transcriptase in that DNA synthesis can occur only by extension of a primer hybridized to an RNA template (25). In the presence of the DNA target, the R2 reverse transcriptase apparently changes its conformation such that either it is able to recognize the R2 transcript or it is able to prime cDNA synthesis without the need for sequence complementarity. Clearly it seems that the biologically significant reaction is the priming of cDNA synthesis by the nicked target site. However, if the R2 transcript does not allow proper positioning of the 3' junction next to the nicked DNA site, then RNA-primed reverse transcription can result. Thus, RNA-primed reverse transcripts should be viewed as aborted attempts to conduct the TPRT reaction. If RNA-primed reverse transcripts are generated *in vivo*, their failure to be integrated would presumably lead to their rapid degradation.

The ability of the R2 reverse transcriptase to use the 3' end of a noncomplementary RNA molecule to prime cDNA synthesis is similar to the properties of the reverse transcriptase

encoded by the Mauriceville plasmid of *N. crassa* (32). Similar to the R2 reverse transcriptase, the Mauriceville enzyme occasionally adds nontemplated nucleotides before its RNA template is engaged (7). The ability of both the Mauriceville plasmid and the R2 reverse transcriptase to use another nucleic acid to prime cDNA synthesis in the absence of any sequence complementarity is a second enzymatic property linking the non-LTR retrotransposable elements to the retroelements found in mitochondrial genomes of fungi. The first, and still most striking, similarity linking the non-LTR retrotransposable elements and retroelements of mitochondria is that both the R2 element and the $\alpha 2$ intron retrotranspose by a TPRT reaction (see reference 8 for a more detailed comparison). On the basis of phylogenetic studies of their reverse transcriptase domains, we originally suggested that the non-LTR retrotransposable elements were more closely related to the Mauriceville plasmid and group II introns of mitochondrial genomes than they were to the LTR retrotransposable elements (37). It is gratifying to now have independent biochemical data suggesting that the reverse transcriptases from these three distinct retroelements have similarities in their priming mechanisms that are entirely distinct from those of the LTR retrotransposable elements and retroviruses. It also raises the interesting possibility that the retrons of bacteria (11, 38), which phylogenetically appear to be most similar to these three non-LTR retroelements, may also share certain of these priming mechanisms. Currently, bacterial retron reverse transcriptases have been shown to prime cDNA synthesis only from an internal 2'-hydroxyl group on the RNA template (17).

ACKNOWLEDGMENTS

This work was supported by a National Institutes of Health grant (GM42790) to T.H.E.

We thank Danna Eickbush, Janet George, and Jin Yang for comments on the manuscript.

REFERENCES

- Aksoy, S., S. Williams, S. Chang, and F. F. Richards. 1990. SLACS retrotransposon from *Trypanosoma brucei gambiense* is similar to mammalian LINEs. *Nucleic Acids Res.* **18**:785-792.
- Boeke, J. D., and V. G. Corces. 1989. Transcription and reverse transcription of retrotransposons. *Annu. Rev. Microbiol.* **43**:403-434.
- Bucheton, A. 1990. I transposable elements and I-R hybrid dysgenesis in *Drosophila*. *Trends Genet.* **6**:16-21.
- Burke, W. D., C. C. Calalang, and T. H. Eickbush. 1987. The site specific ribosomal insertion element type II of *Bombyx mori* (R2Bm) contains the coding sequence for a reverse transcriptase-like enzyme. *Mol. Cell. Biol.* **7**:2221-2230.
- Burke, W. D., D. G. Eickbush, Y. Xiong, J. Jakubczak, and T. H. Eickbush. 1993. Sequence relationship of retrotransposable elements R1 and R2 within and between divergent insect species. *Mol. Biol. Evol.* **10**:163-185.
- Burke, W. D., F. Muller, and T. H. Eickbush. 1995. R4, a non-LTR retrotransposon specific to the large subunit rRNA genes of nematodes. *Nucleic Acids Res.* **23**:4628-4634.
- Chiang, C.-C., J. C. Kennell, L. A. Wanner, and A. M. Lambowitz. 1994. A mitochondrial retroplasmid integrates into mitochondrial DNA by a novel mechanism involving the synthesis of a hybrid cDNA and homologous recombination. *Mol. Cell. Biol.* **14**:6419-6432.
- Curcio, M. J., and M. Belfort. 1996. Retrohoming: cDNA-mediated mobility of group II introns require a catalytic RNA. *Cell* **84**:9-12.
- Eickbush, D. G., and T. H. Eickbush. 1995. Vertical transmission of the retrotransposable elements R1 and R2 during the evolution of the *Drosophila melanogaster* species subgroup. *Genetics* **139**:671-684.
- Eickbush, T. H. 1992. Transposing without ends: the non-LTR retrotransposable elements. *New Biol.* **4**:430-440.
- Eickbush, T. H. 1994. Origin and evolutionary relationships of retroelements, p. 121-157. *In* S. S. Morse (ed.) *The evolutionary biology of viruses*. Raven Press, New York.
- Evans, J. P., and R. D. Palmiter. 1991. Retrotransposition of a mouse L1 element. *Proc. Natl. Acad. Sci. USA* **88**:8792-8795.
- Gabriel, A., and J. D. Boeke. 1991. Reverse transcriptase encoded by a retrotransposon from the trypanosomatid *Crithidia fasciculata*. *Proc. Natl. Acad. Sci. USA* **88**:9794-9798.
- Gabriel, A., T. J. Yen, D. C. Schwartz, C. L. Smith, J. D. Boeke, B. Sollner-Webb, and D. W. Cleveland. 1990. A rapidly rearranging retrotransposon within the minixon gene locus of *Crithidia fasciculata*. *Mol. Cell. Biol.* **10**:615-624.
- Garrett, J. E., D. S. Knutzon, and D. Carroll. 1989. Composite transposable elements in the *Xenopus laevis* genome. *Mol. Cell. Biol.* **9**:3018-3027.
- Hancock, J. M., D. Tautz, and G. A. Dover. 1988. Evolution of the secondary structures and compensatory mutations of the ribosomal RNAs of *Drosophila melanogaster*. *Mol. Biol. Evol.* **5**:393-414.
- Inouye, M., and S. Inouye. 1991. msDNA and bacterial reverse transcriptase. *Annu. Rev. Microbiol.* **45**:163-186.
- Ivanov, V. A., A. A. Melnikov, A. V. Siunov, I. I. Fodor, and Y. V. Illyin. 1991. Authentic reverse transcriptase is coded by jockey, a mobile *Drosophila* element related to mammalian LINEs. *EMBO J.* **10**:2489-2495.
- Jakubczak, J. L., W. D. Burke, and T. H. Eickbush. 1991. Retrotransposable elements R1 and R2 interrupt the rRNA genes of most insects. *Proc. Natl. Acad. Sci. USA* **88**:3295-3299.
- Jensen, S., and T. Heidmann. 1991. An indicator gene for detection of germline retrotransposition in transgenic *Drosophila* demonstrates RNA-mediated transposition of the LINE 1 element. *EMBO J.* **10**:1927-1937.
- Kennell, J. C., J. V. Moran, P. S. Perlman, R. A. Butow, and A. M. Lambowitz. 1993. Reverse transcriptase activity associated with maturase-encoding group II introns in yeast mitochondria. *Cell* **73**:133-146.
- Loh, E. Y., J. F. Elliot, S. Cwirla, L. L. Lanier, and M. D. Davis. 1989. Polymerase chain reaction with single-sided specificity: analysis of T cell receptor δ chain. *Science* **243**:217-220.
- Luan, D. D. Unpublished data.
- Luan, D. D., and T. H. Eickbush. 1995. RNA template requirements for target DNA-primed reverse transcription by the R2 retrotransposable element. *Mol. Cell. Biol.* **15**:3882-3891.
- Luan, D. D., M. H. Korman, J. L. Jakubczak, and T. H. Eickbush. 1993. Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition. *Cell* **72**:595-605.
- Mathias, S. L., A. F. Scott, H. H. Kazazian, J. D. Boeke, and A. Gabriel. 1991. Reverse transcriptase encoded by a human transposable element. *Science* **254**:1808-1810.
- Michel, F., and J.-L. Ferat. 1995. Structure and activities of group II introns. *Annu. Rev. Biochem.* **64**:435-461.
- Pelisson, A., D. J. Finnegan, and A. Bucheton. 1991. Evidence for retrotransposition of the I factor, a LINE element of *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA* **88**:4907-4910.
- Schwarz-Sommer, Z., L. Leclercq, E. Gobel, and H. Saedler. 1987. Cin4, an insert altering the structure of the A1 gene in *Zea mays*, exhibits properties of nonviral retrotransposons. *EMBO J.* **6**:3873-3880.
- Teng, S. C., S. X. Wang, and A. Gabriel. 1995. A new non-LTR retrotransposon provides evidence for multiple distinct site-specific elements in *Crithidia fasciculata* minixon arrays. *Nucleic Acids Res.* **23**:2929-2936.
- Villanueva, M. S., S. P. Williams, C. B. Beard, F. R. Richards, and S. Aksoy. 1991. A new member of a family of site-specific retrotransposons is present in the spliced leader RNA genes of *Trypanosoma cruzi*. *Mol. Cell. Biol.* **11**:6139-6148.
- Wang, H., J. C. Kennell, M. T. R. Kuiper, J. R. Sabourin, R. Saldanha, and A. M. Lambowitz. 1992. The Mauriceville plasmid of *Neurospora crassa*: characterization of a novel reverse transcriptase that begins cDNA synthesis at the 3' end of template RNA. *Mol. Cell. Biol.* **12**:5131-5144.
- Wang, H., and A. M. Lambowitz. 1993. The Mauriceville plasmid reverse transcriptase can initiate cDNA synthesis de novo and may be related to reverse transcriptase and DNA polymerase progenitor. *Cell* **75**:1071-1081.
- Weiner, A. M., P. L. Denninger, and A. Efstratiadis. 1986. Nonviral retrotransposons: genes, pseudogenes, and transposable elements generated by the reverse flow of genetic information. *Annu. Rev. Biochem.* **55**:631-661.
- Whitcomb, J. M., and S. H. Hughes. 1992. Retroviral reverse transcription and integration: process and problems. *Annu. Rev. Cell Biol.* **8**:275-306.
- Xiong, Y., and T. H. Eickbush. 1988. Functional expression of a sequence-specific endonuclease encoded by the retrotransposon R2Bm. *Cell* **55**:235-246.
- Xiong, Y., and T. H. Eickbush. 1988. Similarity of reverse transcriptase-like sequences of viruses, transposable elements, and mitochondrial introns. *Mol. Biol. Evol.* **5**:675-690.
- Xiong, Y., and T. H. Eickbush. 1990. Origin and evolution of retroelements based upon their reverse transcriptase sequences. *EMBO J.* **9**:3351-3362.
- Zimmerly, S., H. Guo, R. Eskes, J. Yang, P. S. Perlman, and A. M. Lambowitz. 1995. A group II intron RNA is a catalytic component of a DNA endonuclease involved in intron mobility. *Cell* **83**:529-538.
- Zimmerly, S., H. Guo, P. S. Perlman, and A. M. Lambowitz. 1995. Group II intron mobility occurs by target DNA-primed reverse transcription. *Cell* **82**:1-10.